

# On Precoding for Constant K-User MIMO Gaussian Interference Channel with Finite Constellation Inputs

Abhinav Ganesan and B. Sundar Rajan

Email: {abhig\_88, bsrajan}@ece.iisc.ernet.in

**Abstract**—This paper considers linear precoding for constant channel-coefficient  $K$ -User MIMO Gaussian Interference Channel (MIMO GIC) where each transmitter- $i$  (Tx- $i$ ), requires to send  $d_i$  independent complex symbols per channel use that take values from fixed finite constellations with uniform distribution, to receiver- $i$  (Rx- $i$ ) for  $i = 1, 2, \dots, K$ . We define the maximum rate achieved by Tx- $i$  using any linear precoder, when the interference channel-coefficients are zero, as the signal to noise ratio (SNR) tends to infinity to be the Constellation Constrained Saturation Capacity (CCSC) for Tx- $i$ . In this paper, we derive a high SNR approximation for the rate achieved by Tx- $i$  when interference is treated as noise, which is given by the mutual information between Tx- $i$  and Rx- $i$ , denoted by  $I[X_i; Y_i]$  where,  $X_i$  denotes the symbols generated at Tx- $i$  before precoding and  $Y_i$  denotes the symbols received at the antennas of Rx- $i$ . Based on this high SNR approximation, we derive a set of necessary and sufficient conditions on the precoders under which  $I[X_i; Y_i]$  tends to CCSC for Tx- $i$ . Interestingly, the precoders designed for interference alignment satisfies these necessary and sufficient conditions. However, finding precoders that align interference is known to be NP-hard in general whereas, the precoders that satisfy the derived necessary and sufficient conditions are easy to find for any given channel-coefficients. Precoding for the case of single antenna at all the nodes correspond to rotation of constellation at the transmitters. As a corollary for the single antenna case we have the result that, using appropriate rotation of constellation at the transmitters, as SNR tends to infinity,  $I[X_i; Y_i]$  tends to CCSC for Tx- $i$  which is independent of the channel-coefficients. This result is in contrast to the Gaussian alphabet case where  $I[X_i; Y_i]$  tends to a value determined by the channel-coefficients, as SNR tends to infinity. Further, we propose a gradient-ascent based algorithm to optimize the sum-rate achieved by precoding with finite constellation inputs and treating interference as noise. Simulation study for a 3-user MIMO GIC with two antennas at each node with  $d_i = 1$  and QPSK inputs, for all  $i$ , shows an improvement of 1.07 bits/sec/Hz in the ergodic sum-rate using the precoders obtained from the proposed algorithm over the precoders that achieve IA, at  $SNR = -2$  dB.

## I. INTRODUCTION

Interference Alignment (IA) has been a focus of intense research on Gaussian interference channels (GICs) in the recent years on the account of the capacity of interference channels being unknown in general and the potential of IA to get close to the sum-capacity for a broad class of interference channels as the signal to noise ratio (SNR) tends to infinity. The scaling of the sum-capacity with  $\log SNR$  is known as the sum degrees of freedom (DoF) of the GIC [1]. The sum-DoF is known to be  $K/2$ , with probability 1, for the  $K$ -user

GIC [1] when all the nodes are equipped with a single antenna and time-varying channel gains are assumed. The result was proved by using linear precoding at the transmitters over an arbitrarily large number of symbol extensions and zero forcing at the receivers. In a later work [2], the sum-DoF is shown to be  $K/2$  with probability 1 even in the case of constant channel-coefficients (that are drawn from a continuous distribution), i.e., the channel gains do not vary with time, using a non-linear IA technique. When all the nodes are equipped with  $M$  antennas, the sum-DoF was shown to be  $3M/2$  for the 3-user GIC [1]. This result was proved by linear precoding over the transmit antennas without the use of symbol extensions and holds true even in the case of constant channel coefficients. Later [3], with the assumption of constant channel coefficients and using a non-linear IA technique, the sum-DoF when  $K \geq \frac{M+N}{\gcd(M,N)}$  was found to be equal to  $\frac{MN}{M+N}K$  where,  $\gcd(M, N)$  denotes the greatest common divisor of  $M$  and  $N$ , with  $M$  being the number of antennas at each transmitter and  $N$  being the number of receive antennas at each receiver. All the works cited above assumed full channel state information at all the transmitters (CSIT) and receivers (CSIR). The notion of sum-DoF involves scaling of sum-rate as  $\log SNR$  at high SNR and therefore, Gaussian input alphabets or lattice codes are always used in the study of sum-DoF. However, in all practical scenarios finite constellations like  $M$ -QAM and  $M$ -PSK are used at the inputs. *With the constraint of finite constellation inputs, it is not known whether IA is optimal in some sense.*

Linear precoding for optimizing the mutual information between the input and the output has been studied for the single user MIMO channel with finite constellation inputs in [4]- [6]. Constellation rotation for optimizing the sum-capacity for SISO Multiple Access Channel (MAC) with finite constellation inputs has been examined in [7] and linear precoding for weighted sum-rate maximization in MIMO MAC with finite constellation inputs has been studied in [8]. Note that linear precoding for the SISO MAC corresponds to constellation rotation at the transmitter.

Recently, there has been some progress on the analysis of finite constellation effects in 2-user SISO GIC [9], [10]. In [9], constellation rotation was found to increase the constellation constrained sum-capacity of 2-user SISO Gaussian strong interference channel [11], [12], and in [10], a metric to find the optimum angle of rotation was proposed. In this paper,

we examine achievable rate-tuples with linear precoding for  $K$ -user MIMO Gaussian Interference Channel (GIC) with finite constellation inputs. Specifically, we treat interference as noise, i.e., each transmitter reveals its codebook only to its intended receiver. The maximum rate achievable under such a circumstance for transmitter- $i$  (Tx- $i$ ) is given by mutual information between the input generated by Tx- $i$  and the output at receiver Rx- $i$ . The channel conditions and values of SNR under which the decoding scheme of treating interference as noise with Gaussian alphabet inputs is sum-capacity optimal was found for the 2-user SISO GIC in [13]- [15], for the  $K$ -user SISO GIC in [16] and for the 2-user MIMO GIC in [17], [18]. For given values of channel gains, with Gaussian input alphabets, as the SNR tends to infinity, treating interference as noise is not sum-capacity optimal [19].

*With the constraint of finite constellation inputs, it is not clear whether treating interference as noise is optimal in some sense.* First, we need to define a notion of optimality under the constraint of fixed finite constellation inputs and then analyse decoding and transmit schemes with that notion of optimality. Consider a scenario where each transmitter- $i$  (Tx- $i$ ), requires to send  $d_i$  independent complex symbols per channel use that take values from fixed finite constellations with uniform distribution, for  $i = 1, 2, \dots, K$ , to receiver- $i$  (Rx- $i$ ). Throughout this paper, we assume that none of the direct channel gains are zero. For a  $K$ -user MIMO GIC with finite constellation inputs, as a measure of optimality of linear precoding in the high SNR sense, we introduce the notion of *Constellation Constrained Saturation Capacity* (CCSC) which is defined as follows.

*Definition 1:* The maximum rate achieved by Tx- $i$  as SNR tends to infinity, using any linear precoder, when the interference channel-coefficients are zero is termed as the Constellation Constrained Saturation Capacity (CCSC) for Tx- $i$ .

For the ease of exposition, throughout the paper, we assume that the constellations used for the symbols are all the same at all the transmitters, and is of cardinality  $M$ . Hence, the CCSC for Tx- $i$  is given by  $\log_2 M^{d_i}$ .

In this paper, with the assumption of constant  $K$ -user MIMO GIC with full global knowledge of channels gains, and finite constellation inputs, we derive a set of necessary and sufficient conditions on the precoders under which treating interference as noise at Rx- $i$  will achieve a rate for Tx- $i$  that tends to CCSC for Tx- $i$ , for all  $i$ , as SNR tends to infinity. Precoders satisfying these necessary and sufficient conditions exist for all direct and cross channel gains, and are termed as *CCSC optimal precoders*. Hence, in the case of finite constellation inputs with the use of appropriate precoders, the rate tuples obtained by treating interference as noise tend to values that are independent of the channel gains. For a  $K$ -user SISO GIC, this result is in contrast with the Gaussian input alphabet case where the rate tuples obtained by treating interference as noise tend to values dictated by the channel gains, as the SNR tends to infinity. Interestingly, the precoders that achieve IA, if feasible, are also CCSC optimal precoders. However, finding precoders that align interference is known

to be NP-hard [20] in general whereas, the precoders that satisfy the derived necessary and sufficient conditions are easy to find for any given channel-coefficients. Since finite SNR is of more practical interest, we propose a gradient-ascent based algorithm to optimize the precoders for the sum-rate achieved by treating interference as noise. The contributions of the paper are summarized below.

- For a constant  $K$ -user MIMO GIC using finite constellation inputs with precoding, a high SNR approximation for the rate tuples achieved by treating interference as noise at the receivers is derived (see Theorem 1 in Section III). Based on this approximation, we derive a set of necessary and sufficient conditions under which the precoders are CCSC optimal (see Theorem 2 Section III). These conditions are satisfied with probability 1 when the entries of the precoders are chosen from any continuous distribution. It is observed that the precoders that achieve IA, if feasible, are CCSC optimal.
- For the finite SNR case, we propose a gradient-ascent based algorithm to improve the sum-rate achieved by treating interference as noise using finite constellation inputs with precoding. Simulation studies indicate considerable improvement in the ergodic sum-rate with precoders obtained using the proposed algorithm for a 3-user MIMO GIC with 2 antennas at all the nodes,  $d_i = 1$ , for all  $i$ , and QPSK inputs, at low and moderate SNRs over that obtained using the IA solution of [1].

The paper is organized as follows. The system model is formally introduced in Section II. In Section III, a set of necessary and sufficient conditions for CCSC optimal precoders is derived. In Section IV, our gradient-ascent based algorithm for optimizing the sum-rate using precoders and treating interference as noise at the receivers is given, and a simulation result comparing its performance with respect to that of precoders satisfying IA is presented. Section V concludes the paper.

*Notations:* For a random variable  $X$  which takes value from the set  $\mathcal{X}$ , we assume some ordering of its elements and use  $x^i$  to represent the  $i$ -th element of  $\mathcal{X}$ . Realization of the random variable  $X$  is denoted as  $x$ . The notation  $\text{diag}(V_1, V_2, \dots, V_n)$  denotes a block diagonal matrix formed by the matrices  $V_i$ ,  $i = 1, 2, \dots, n$ . The  $i^{\text{th}}$  coordinate of a complex vector  $X$  is denoted by  $X(i)$ . The 2-norm of a complex vector  $X$  is denoted by  $\|X\|$ . The cardinality of a set  $\mathcal{X}$  is denoted by  $|\mathcal{X}|$ . For a complex number  $a$ ,  $\Re\{a\}$  and  $\Im\{a\}$  denote the real and imaginary parts of  $a$  respectively. For two complex numbers  $a$  and  $b$ , the notation  $a > b$  denotes that  $|\Re\{a\}| > |\Re\{b\}|$  and  $|\Im\{a\}| > |\Im\{b\}|$ . The notation  $\underline{0}$  represents the zero vector whose size will be clear from the context. All the logarithms in the paper are to the base 2.

## II. SYSTEM MODEL

Tx- $i$  intends to communicate with Rx- $i$ , for  $i = 1, 2, \dots, K$ , as shown in Fig. 1. Without loss of generality, let  $P$  denote the

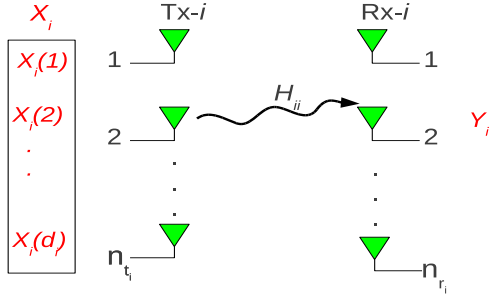


Fig. 1. System Model.

power constraint at all the transmitters. The signal received at Rx- $j$  is given by

$$Y_j = \sum_{i=1}^K \sqrt{P} H_{ij} V_i X_i + N_j$$

where,  $H_{ij}$  denotes the constant channel matrix from Tx- $i$  to Rx- $j$ ,  $V_i$  denotes the precoder at Tx- $i$ ,  $X_i$  denotes the complex symbol vector generated at Tx- $i$ ,  $N_j$  denotes the noise random vector whose coordinates represent independent and identically distributed zero mean unit variance circularly symmetric complex Gaussian random variables. The sizes of the matrices  $H_{ij}$ ,  $V_i$ ,  $X_i$ , and  $N_j$  are given by  $n_{r_j} \times n_{t_i}$ ,  $n_{t_i} \times d_i$ ,  $d_i \times 1$ , and  $n_{r_j} \times 1$  respectively, where  $n_{r_j}$  and  $n_{t_i}$  denote the number of receive and transmit antennas at Rx- $j$  and Tx- $i$  respectively, and  $d_i$  denotes the number of independent complex symbols per channel use that Tx- $i$  wants to transmit to Rx- $i$ . These complex symbols are assumed to take values from finite constellations with uniform distribution over its elements. For simplicity of exposition, we assume that the finite constellations used are all the same at all the transmitters and are of cardinality  $M$ . The results of this paper apply with simple modifications when this is not the case. The finite constellation used is denoted by  $\mathcal{S}$  and is of unit power.

### III. CCSC OPTIMAL PRECODERS

In this section, we shall derive a set of necessary and sufficient on the precoders for CCSC optimality which is taken to be a measure of optimality for linear precoding in the high SNR regime for the finite constellation input case. Rate achievable for Tx- $i$  by treating interference as noise at Rx- $i$ , for all  $i$ , is given by  $R_i < I[X_i; Y_i]$ . Our focus will be on the boundary point given by  $I[X_i; Y_i]$ , for all  $i$ . Let  $X = [X_1 \ X_2 \ \cdots \ X_K]^T$ . The effective channel matrix from all the transmitters to Rx- $i$  is given by  $H_i = [H_{1i} \ H_{2i} \ \cdots \ H_{Ki}]$ . Define

$$V = \text{diag}(V_1, V_2, \cdots, V_K).$$

Using the chain rule for mutual information [21],

$$I[X_i; Y_i] = I[X_1, X_2, \cdots, X_K; Y_i] - I[X_1, X_2, \cdots, X_K; Y_i | X_i].$$

With uniform distribution assumed over the elements of the constellation, the expression for  $I[X_1, X_2, \cdots, X_K; Y_i]$  is given by (2) (at the top of the next page) which is derived in a similar way as in [10]. Define the matrix

$$A^{k_1, k_2} = H_i V (x^{k_1} - x^{k_2}).$$

The following theorem gives a high SNR approximation for  $I[X_1, X_2, \cdots, X_K; Y_i]$ .

*Theorem 1:* At high  $P$ , the mutual information  $I[X_1, X_2, \cdots, X_K; Y_i]$  can be approximated by

$$\log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i} - 1} \left[ \log \left( \sum_{k_2=0}^{M^{\sum_{i=1}^K d_i} - 1} e^{-\|\sqrt{P} A^{k_1, k_2}\|^2} \right) \right]. \quad (1)$$

*Proof:* Define the sets

$$\mathcal{A}^{k_1} = \{k_2 \neq k_1 \mid A^{k_1, k_2} = 0\}, \quad (10)$$

$$\mathcal{A}^{k_1, k_2} = \{l \mid A^{k_1, k_2}(l) = 0\}. \quad (11)$$

The expression in (2) is now re-written as in (3) and (4). At high values of  $P$ , the following inequality holds good for  $k_2 \neq k_1$ ,  $k_2 \notin \mathcal{A}^{k_1}$ , and for all  $l \notin \mathcal{A}^{k_1, k_2}$ .

$$|\Re\{\sqrt{P} A^{k_1, k_2}(l)\}| >> \frac{3}{\sqrt{2}} \text{ and } |\Im\{\sqrt{P} A^{k_1, k_2}(l)\}| >> \frac{3}{\sqrt{2}}. \quad (12)$$

Now, note that for  $k_2 \neq k_1$ ,  $k_2 \notin \mathcal{A}^{k_1}$ , and for all  $l \notin \mathcal{A}^{k_1, k_2}$ ,  $\sqrt{P} A^{k_1, k_2}(l) >> n_j(l)$  when  $|n_{jR}(l)| \leq 3$  and  $|n_{jI}(l)| \leq 3$  where,  $n_{jR}(l)$  and  $n_{jI}(l)$  represent the real and imaginary parts of  $n_j(l)$ . The value of  $n_j(l)$  becomes comparable to  $\sqrt{P} A^{k_1, k_2}(l)$  only if either  $|n_{jR}(l)| > 3$  or  $|n_{jI}(l)| > 3$ . However, the probability of such an event occurring is extremely small because the variances of the Gaussian random variables  $n_{jR}(l)$  and  $n_{jI}(l)$  are equal to  $\frac{1}{2}$ . Hence, the contribution of such an event to the integral in (5) is very small that it can be neglected at high  $P$ . Therefore, on account of (12), the approximations in (6) and (7) are valid. The equation (8) follows from the fact that probability distribution integrates to 1, (9) follows from (11), and the proposed approximation in (1) is obtained directly by re-writing (9). ■

*Remark 1:* Note that we cannot straightforwardly argue that at high powers  $P$ , the noise  $N_i$  in (2) can be neglected. This is because the value of  $\|N_i\|$  can be of the order of  $\|A^{k_1, k_2}\|$ . The proof uses the fact that the probability of such an event is very small and hence, can be neglected. The greater the power, the better is the approximation. A similar approximation was developed for the SISO case in [10]. The approximation given in Theorem 1 is a generalization of the approximation derived in [10].

Let  $X_{\neq} = [X_1 \ X_2 \ \cdots \ X_{i-1} \ X_{i+1} \ \cdots \ X_K]^T$ . The channel matrix from all the transmitters, with the exclusion of Tx- $i$ , to Rx- $i$  is given by

$$H_{\neq} = [H_{11} \ H_{12} \ \cdots \ H_{i-1, i} \ H_{i+1, i} \ \cdots \ H_{Ki}].$$

$$I(X_1, X_2, \dots, X_K; Y_i) = \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} E_{N_i} \left[ \log \sum_{k_2=0}^{M^{\sum_{i=1}^K d_i-1}} \exp - \left( \|N_i + \sqrt{P} H_i V(x^{k_1} - x^{k_2})\|^2 - \|N_i\|^2 \right) \right] \quad (2)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} E_{N_i} \left[ \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} \exp - \left( \|N_i + \sqrt{P} A^{k_1, k_2}\|^2 - \|N_i\|^2 \right) \right) \right] \quad (3)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} E_{N_j} \left[ \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} \exp - \left( \sum_{l \notin \mathcal{A}^{k_1, k_2}} \left( |N_i(l) + \sqrt{P} A^{k_1, k_2}(l)|^2 - |N_i(l)|^2 \right) \right) \right) \right] \quad (4)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \int_{\substack{(n_{iR}(l), n_{iI}(l)) \\ l \notin \mathcal{A}^{k_1, k_2}}} \prod_{l \notin \mathcal{A}^{k_1, k_2}} p_{N_i(l)}(n_j(l)) \right. \\ \left. \times \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-\left( \sum_{l \notin \mathcal{A}^{k_1, k_2}} \left( |n_i(l) + \sqrt{P} A^{k_1, k_2}(l)|^2 - |n_i(l)|^2 \right) \right)} \right) \prod_{l \notin \mathcal{A}^{k_1, k_2}} d n_i(l) \right] \quad (5)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \int_{\substack{(n_{iR}(l), n_{iI}(l)) \\ l \notin \mathcal{A}^{k_1, k_2}}} \prod_{l \notin \mathcal{A}^{k_1, k_2}} p_{N_i(l)}(n_i(l)) \right. \\ \left. \times \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-\left( \sum_{l \notin \mathcal{A}^{k_1, k_2}} \left( |\sqrt{P} A^{k_1, k_2}(l)|^2 - |n_i(l)|^2 \right) \right)} \right) \prod_{l \notin \mathcal{A}^{k_1, k_2}} d n_i(l) \right] \quad (6)$$

$$\approx \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \int_{\substack{(n_{jR}(l), n_{jI}(l)) \\ l \notin \mathcal{A}^{k_1, k_2}}} \prod_{l \notin \mathcal{A}^{k_1, k_2}} p_{N_j(l)}(n_j(l)) \right. \\ \left. \times \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-\left( \sum_{l \notin \mathcal{A}^{k_1, k_2}} \left( |\sqrt{P} A^{k_1, k_2}(l)|^2 \right) \right)} \right) \prod_{l \notin \mathcal{A}^{k_1, k_2}} d n_j(l) \right] \quad (7)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-\left( \sum_{l \notin \mathcal{A}^{k_1, k_2}} \left( |\sqrt{P} A^{k_1, k_2}(l)|^2 \right) \right)} \right) \right] \quad (8)$$

$$= \log M^{\sum_{i=1}^K d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-\left( \sum_{l=1}^{n_i} \left( |\sqrt{P} A^{k_1, k_2}(l)|^2 \right) \right)} \right) \right] \quad (9)$$

Define the matrices

$$V_{\check{}} = \text{diag}(V_1, V_2, \dots, V_{i-1}, V_{i+1}, \dots, V_K).$$

$$B^{i_1, i_2} = H_{\check{}} V_{\check{}} \begin{pmatrix} x_{\check{}}^{i_1} - x_{\check{}}^{i_2} \end{pmatrix},$$

for  $i_1, i_2 = 0, 1, \dots, M^{\sum_{j \neq i} d_j} - 1$ . Similar to (1), we have the following approximation for  $I[X_1, X_2, \dots, X_K; Y_i | X_i]$  at high  $P$ .

$$\log M^{\sum_{j \neq i}^K d_j} - \frac{1}{M^{\sum_{j \neq i} d_j}} \sum_{i_1=0}^{M^{\sum_{j \neq i} d_j-1}} \left[ \log \left( \sum_{i_2=0}^{M^{\sum_{j \neq i} d_j-1}} e^{-\| \sqrt{P} B^{i_1, i_2} \|^2} \right) \right] \quad (13)$$

Hence, a high SNR approximation for  $I[X_i; Y_i]$  is given by

$$I[X_i; Y_i] \approx \log M^{d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^K d_i-1}} \left[ \log \left( \sum_{k_2=0}^{M^{\sum_{i=1}^K d_i-1}} e^{-\| \sqrt{P} A^{k_1, k_2} \|^2} \right) \right] \\ + \frac{1}{M^{\sum_{j \neq i} d_j}} \sum_{i_1=0}^{M^{\sum_{j \neq i} d_j-1}} \left[ \log \left( \sum_{i_2=0}^{M^{\sum_{j \neq i} d_j-1}} e^{-\| \sqrt{P} B^{i_1, i_2} \|^2} \right) \right]. \quad (14)$$

Now, define the set

$$\mathcal{B}^{i_1} = \{i_2 \neq i_1 \mid B^{i_1, i_2} = 0\}. \quad (15)$$

The following theorem gives a set of necessary and sufficient conditions under which the above approximation tends to  $\log M^{d_i}$  as  $P$  tends to infinity and hence, gives a set of

necessary and sufficient conditions under which the precoders are CCSC optimal.

**Theorem 2:** The approximation for  $I[X_i; Y_i]$  given in (14) tends to  $\log M^{d_i}$  as  $P$  tends to infinity iff

$$H_{ii}V_i(x_i^{p_{i1}} - x_i^{p_{i2}}) + \sum_{\substack{k \neq i \\ k=1}}^K H_{kj}V_k(x_k^{p_{k1}} - x_k^{p_{k2}}) \neq \underline{0}, \quad (16)$$

$$\forall p_{i1} \neq p_{i2}, \forall p_{k1}, \forall p_{k2}$$

where,  $p_{k1}, p_{k2} = 0, 1, \dots, M^{d_i} - 1$ .

*Proof:* The summation-term of the second term in (14) is re-written as

$$\begin{aligned} & \sum_{k_1=0}^{M^{\sum_{i=1}^{d_i} d_i - 1}} \left[ \log \left( 1 + |\mathcal{A}^{k_1}| + \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-P} \left( \|A^{k_1, k_2}\|^2 \right) \right) \right] \\ &= \sum_{k_1=0}^{M^{\sum_{i=1}^{d_i} d_i - 1}} \log \left( 1 + |\mathcal{A}^{k_1}| \right) + \\ & \quad \log \left( 1 + \frac{1}{1 + |\mathcal{A}^{k_1}|} \sum_{\substack{k_2 \neq k_1 \\ k_2 \notin \mathcal{A}^{k_1}}} e^{-P} \left( \|A^{k_1, k_2}\|^2 \right) \right) \\ & \xrightarrow{P \rightarrow \infty} \sum_{k_1=0}^{M^{\sum_{i=1}^{d_i} d_i - 1}} \log \left( 1 + |\mathcal{A}^{k_1}| \right). \end{aligned}$$

Similarly, the summation-term of the last term in (14) tends to  $\sum_{i_1=0}^{M^{\sum_{j \neq i} d_j}} \log(1 + |\mathcal{B}^{i_1}|)$  as  $P$  tends to infinity. Hence, as  $P$  tends to infinity, (14) tends to

$$\begin{aligned} \log M^{d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{k_1=0}^{M^{\sum_{i=1}^{d_i} d_i - 1}} \log(1 + |\mathcal{A}^{k_1}|) \\ + \frac{1}{M^{\sum_{j \neq i} d_j}} \sum_{i_1=0}^{M^{\sum_{j \neq i} d_j}} \log(1 + |\mathcal{B}^{i_1}|). \end{aligned} \quad (17)$$

Now, define the sets

$$\begin{aligned} \mathcal{C}^{p_{11}, \dots, p_{K1}} &= \{ (p_{12}, p_{22}, \dots, p_{K2}) \neq (p_{11}, p_{21}, \dots, p_{K1}) \mid \\ & H_{ii}V_i(x_i^{p_{i1}} - x_i^{p_{i2}}) + \sum_{\substack{k \neq i \\ k=1}}^K H_{kj}V_k(x_k^{p_{k1}} - x_k^{p_{k2}}) \neq \underline{0} \}, \\ \mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}} &= \{ (p_{12}, \dots, p_{i-1,2}, p_{i+1,2}, \dots, p_{K2}) \\ & \neq (p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}) \mid \\ & \sum_{\substack{k \neq i \\ k=1}}^K H_{kj}V_k(x_k^{p_{k1}} - x_k^{p_{k2}}) \neq \underline{0} \} \end{aligned}$$

where,  $p_{l1}, p_{l2} = 0, 1, \dots, M^{d_l} - 1$ . Observe that the set of all  $\mathcal{C}^{p_{11}, \dots, p_{K1}}$  has a one-one correspondence with the set of all  $\mathcal{A}^{k_1}$ , and the set of all  $\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}$  has a one-one correspondence with the set of all  $\mathcal{B}^{i_1}$ . Hence, (17) can be re-written as

$$\begin{aligned} \log M^{d_i} - \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{C}^{p_{11}, \dots, p_{K1}}|) \\ + \frac{1}{M^{\sum_{j \neq i} d_j}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{\substack{p_{i-1,1}=0 \\ p_{i+1,1}=0}}^{M^{d_{i-1}-1} M^{d_{i+1}-1}} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}|). \end{aligned} \quad (18)$$

Now, the set  $\mathcal{C}^{p_{11}, \dots, p_{K1}}$  can be written as a disjoint union of two sets, i.e.,

$$\mathcal{C}^{p_{11}, \dots, p_{K1}} = \mathcal{C}_1^{p_{11}, \dots, p_{K1}} \cup \mathcal{C}_2^{p_{11}, \dots, p_{K1}} \quad (19)$$

where,  $\mathcal{C}_1^{p_{11}, \dots, p_{K1}}$  and  $\mathcal{C}_2^{p_{11}, \dots, p_{K1}}$  are defined in (20) and (22) respectively (given at the top of the next page). The set  $\mathcal{C}_1^{p_{11}, \dots, p_{K1}}$  can be re-defined as in (21). Since  $\sum_{k \neq i}^K H_{kj}V_k(x_k^{p_{k1}} - x_k^{p_{k2}})$  is independent of the indices  $p_{i1}$  and  $p_{i2}$ , the set  $\mathcal{C}_1^{p_{11}, \dots, p_{K1}}$  is the same for all  $p_{i1} = 0, 1, \dots, M^{d_i} - 1$ . Now, note that the set of all  $\mathcal{C}_1^{p_{11}, \dots, p_{i-1,1}, 0, p_{i+1,1}, \dots, p_{K1}}$  has a one-one correspondence with the set of all  $\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}$  which follows from the definitions of the respective sets. Hence, (18) can be re-written as

$$\begin{aligned} \log M^{d_i} &- \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{\substack{p_{i-1,1}=0 \\ p_{i+1,1}=0}}^{M^{d_{i-1}-1} M^{d_{i+1}-1}} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}| + |\mathcal{C}_2^{p_{11}, \dots, p_{K1}}|) \\ &+ \frac{1}{M^{\sum_{j \neq i} d_j}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{\substack{p_{i-1,1}=0 \\ p_{i+1,1}=0}}^{M^{d_{i-1}-1} M^{d_{i+1}-1}} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}|) \\ &= \log M^{d_i} \\ &- \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{\substack{p_{i-1,1}=0 \\ p_{i+1,1}=0}}^{M^{d_{i-1}-1} M^{d_{i+1}-1}} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}| + |\mathcal{C}_2^{p_{11}, \dots, p_{K1}}|) \\ &+ \frac{1}{M^{\sum_{i=1}^K d_i}} \sum_{p_{11}=0}^{M^{d_1}-1} \dots \sum_{\substack{p_{i-1,1}=0 \\ p_{i+1,1}=0}}^{M^{d_{i-1}-1} M^{d_{i+1}-1}} \dots \sum_{p_{K1}=0}^{M^{d_K}-1} \log(1 + |\mathcal{D}^{p_{11}, \dots, p_{i-1,1}, p_{i+1,1}, \dots, p_{K1}}|). \end{aligned}$$

Clearly, if  $|\mathcal{C}_2^{p_{11}, \dots, p_{K1}}| > 1$  for some  $(p_{11}, \dots, p_{K1})$  then, the second term in the above equation is strictly greater than the last term and hence, as  $P$  tends to infinity,  $I[X_i; Y_i]$  tends to a value that is strictly less than  $\log M^{d_i}$ . If  $|\mathcal{C}_2^{p_{11}, \dots, p_{K1}}| = 0$  for all  $p_{l1}$  then, the second term in the above equation is equal to the last term and hence,  $I[X_i; Y_i]$  tends to a value equal to  $\log M^{d_i}$ . Thus,  $I[X_i; Y_i]$  tends to  $\log M^{d_i}$  as  $P$  tends to infinity iff (16) is satisfied. ■

**Remark 2:** The result of Theorem 2 means that the rate achieved by treating interference as noise at high  $P$  tends to CCSC for Tx- $i$  iff, in the absence of the Gaussian noise, two different symbol vectors  $x_i^{p_{i1}}$  and  $x_i^{p_{i2}}$  sent by Tx- $i$  should not

$$\mathcal{C}_1^{p_{11}, \dots, p_{K1}} = \{ (p_{12}, p_{22}, \dots, p_{K2}) \neq (p_{11}, p_{21}, \dots, p_{K1}) \text{ for } p_{i2} = p_{i1} \mid H_{ii} V_i (x_i^{p_{i1}} - x_i^{p_{i2}}) + \sum_{\substack{k \neq i \\ k=1}}^K H_{kj} V_k (x_k^{p_{k1}} - x_k^{p_{k2}}) = \underline{0} \}, \quad (20)$$

$$\Rightarrow \mathcal{C}_1^{p_{11}, \dots, p_{K1}} = \{ (p_{12}, p_{22}, \dots, p_{K2}) \neq (p_{11}, p_{21}, \dots, p_{K1}) \text{ for } p_{i2} = p_{i1} \mid \sum_{\substack{k \neq i \\ k=1}}^K H_{kj} V_k (x_k^{p_{k1}} - x_k^{p_{k2}}) = \underline{0} \} \quad (21)$$

$$\mathcal{C}_2^{p_{11}, \dots, p_{K1}} = \{ (p_{12}, p_{22}, \dots, p_{K2}) \neq (p_{11}, p_{21}, \dots, p_{K1}) \text{ for } p_{i2} \neq p_{i1} \mid H_{ii} V_i (x_i^{p_{i1}} - x_i^{p_{i2}}) + \sum_{\substack{k \neq i \\ k=1}}^K H_{kj} V_k (x_k^{p_{k1}} - x_k^{p_{k2}}) = \underline{0} \} \quad (22)$$

map to the same symbol vector at Rx- $i$  for any data symbol transmitted by the interfering transmitters.

*Remark 3:* For a given value of channel gains with none of the direct channel gains being 0, when the entries of the precoders are chosen from any continuous distribution (say, standard normal distribution) the probability of the event

$$H_{ii} V_i (x_i^{p_{i1}} - x_i^{p_{i2}}) + \sum_{\substack{k \neq i \\ k=1}}^K H_{kj} V_k (x_k^{p_{k1}} - x_k^{p_{k2}}) = \underline{0},$$

to occur for any  $p_{i1} \neq p_{i2}$  and for any  $(p_{k1}, p_{k2})$  is zero. By appropriate scaling of the precoders thus obtained, with probability 1, we have CCSC optimal precoders.

*Remark 4:* Interference alignment, if feasible [22] for the given values of  $n_{t_i}$ ,  $n_{r_i}$ , and  $d_i$  involves finding precoders such that the signal sub-space at Rx- $i$ , generated by  $[H_{ii} V_i]$ , is linearly independent of the interference sub-space, generated by  $[H_{1i} V_1 \cdots H_{i-1,i} V_{i-1} \ H_{i+1,i} V_{i+1} \cdots H_{K,i} V_K]$ , and the matrix  $[H_{ii} V_i]$  is full-rank [1]. The CCSC optimality condition in (16) can be rewritten as

$$\begin{bmatrix} H_{ii} V_i & H_{1i} V_1 & \cdots & H_{i-1,i} V_{i-1} & H_{i+1,i} V_{i+1} & \cdots & H_{K,i} V_K \end{bmatrix} \times \begin{bmatrix} (x_i^{p_{i1}} - x_i^{p_{i2}}) \\ (x_1^{p_{11}} - x_1^{p_{12}}) \\ \vdots \\ (x_{i-1}^{p_{i-1,1}} - x_{i-1}^{p_{i-1,2}}) \\ (x_{i+1}^{p_{i+1,1}} - x_{i+1}^{p_{i+1,2}}) \\ \vdots \\ (x_K^{p_{K1}} - x_K^{p_{K2}}) \end{bmatrix} \neq \underline{0}, \forall p_{i1} \neq p_{i2}.$$

Since, with precoders that achieve IA, the signal sub-space at Rx- $i$  is linearly independent of the interference sub-space and  $[H_{ii} V_i]$  is full-rank, the above condition is satisfied for all  $i$ . Hence, IA precoders are also CCSC optimal precoders. However, in general, finding such precoders are NP-hard [20] whereas finding CCSC optimal precoders are easy to find as explained in the previous remark.

*Example 1:* Consider a MIMO GIC with  $K = 3$ ,  $n_{t_i} = n_{r_i} = 1$ ,  $d_i = 1$  for all  $i$ , and the finite constellation used is QPSK. The channel matrix and the precoders are given by

$$H = \begin{bmatrix} -0.9 + 0.4i & -1.7 - 1.40i & 1.5 + 5.0i \\ 2.6 - 0.9i & -0.9 - 2.8i & 0.04 + 0.88i \\ -2.9 - 5.2i & -10.2 + 0.7i & -0.5 + 2.4i \end{bmatrix},$$

$$V_1 = 1, V_2 = e^{i\frac{\pi}{3}}, V_3 = 1$$

where, the matrix element  $[H]_{ij}$  represents the channel gain from Rx- $i$  to Tx- $j$ . The mutual information  $I[X_i; Y_i]$  evaluated

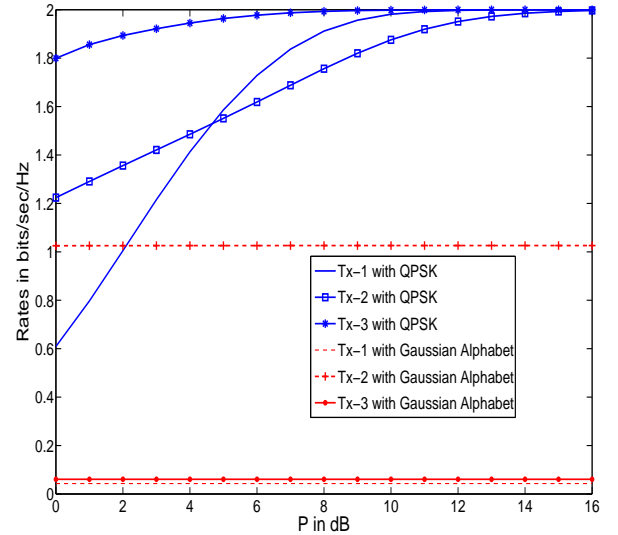


Fig. 2. Rates in bits/sec/Hz vs  $P$  in dB for Example 1.

using Monte-Carlo simulation is plotted for QPSK inputs and Gaussian inputs, for all  $i$ , in Fig. 2. The chosen precoders satisfy (16) and hence,  $I[X_i; Y_i]$  saturates to 2 bits/sec/Hz for all  $i$ , as  $P$  tends to infinity in the QPSK case whereas, for the Gaussian alphabet case, the saturation rate is determined by the channel gains. The saturation value of  $I[X_i; Y_i]$  in the Gaussian alphabet case is given by  $\log \left( 1 + \frac{|h_{ii}|^2}{\sum_{k \neq i} |h_{ki}|^2} \right)$  which evaluates to 0.04, 1.02, and 0.06 bits/sec/Hz for Tx-1, Tx-2 and Tx-3 respectively in this example.

*Example 2:* Consider a MIMO GIC with  $K = 3$ ,  $n_{t_i} = n_{r_i} = 2$ ,  $d_i = 1$  for all  $i$ , and the finite constellation used is QPSK. Let the effective matrix from all the transmitters to all the receivers be given by

$$H = \begin{bmatrix} H_{11} & H_{21} & H_{31} \\ H_{12} & H_{22} & H_{32} \\ H_{13} & H_{23} & H_{33} \end{bmatrix}$$

where,  $H_{ij}$  is the  $2 \times 2$  channel matrix from Tx- $i$  to Rx- $j$ . The precoders and the channel matrices are given in (23) and (24) respectively (given at the top of the next page). The QPSK points are given by  $\left( \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i, \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i \right)$ . Now, (16)

$$V_1 = \begin{bmatrix} 0.66 + 0.74i \\ 0.13 + 0.99i \end{bmatrix}, V_2 = \begin{bmatrix} 0.9883 + 0.1524i \\ 0.4538 + 0.8911i \end{bmatrix}, V_3 = \begin{bmatrix} 0.7044 + 0.7098i \\ 0.1603 + 0.9871i \end{bmatrix}, \quad (23)$$

$$H = \begin{bmatrix} 0.5756 - 0.0565i & 0.7524 - 0.1375i & 0.1697 - 0.1069i & 0.0124 - 0.2002i & 0 & 0 \\ 0.1610 + 0.3766i & -0.0010 + 0.2005i & 0.8758 - 0.0689i & -0.1285 + 0.0605i & 0 & 0 \\ -1.1533 - 0.1280i & -0.6361 + 1.4658i & -1.3069 + 0.1090i & 0.0427 + 0.2488i & -0.0028 + 0.2215i & -1.0597 - 0.2708i \\ -1.7763 - 0.3748i & 0.5341 + 0.0966i & -0.9491 + 0.8074i & -1.0773 - 1.7202i & 0.9616 - 1.2130i & -0.6077 + 0.6970i \\ -1.7082 - 0.4948i & -0.6101 - 0.4739i & -0.2226 - 4.2486i & -0.8216 + 0.4808i & 0.9572 + 1.8870i & -1.4428 - 1.4353i \\ -1.3014 - 0.5614i & 1.2515 + 0.3414i & 0.4242 + 0.0202i & 0.0138 - 0.8740i & 0.3393 - 1.3451i & 0.9498 - 1.0932i \end{bmatrix}, \quad (24)$$

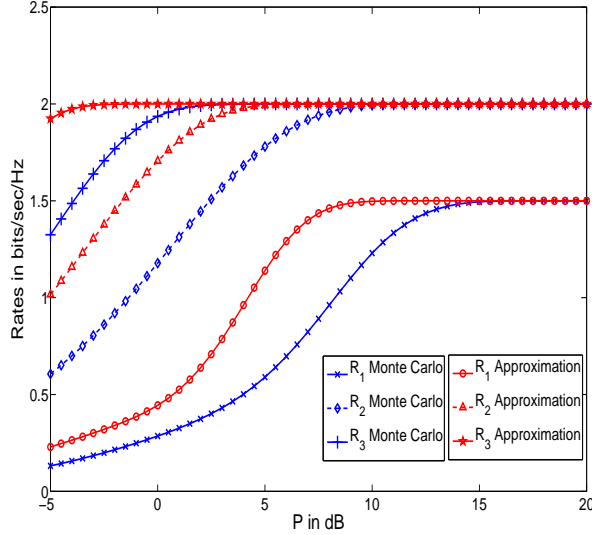


Fig. 3. Rates in bits/sec/Hz vs  $P$  in dB for Example 1.

is not satisfied for  $i = 1$  because

$$[H_{11}V_1 \ H_{21}V_2 \ H_{31}V_3] \times \begin{bmatrix} (x_1^{p_{11}} - x_1^{p_{12}}) \\ (x_2^{p_{21}} - x_2^{p_{22}}) \\ (x_3^{p_{31}} - x_3^{p_{32}}) \end{bmatrix} = \underline{0}, \text{ for } \begin{bmatrix} (x_1^{p_{11}} - x_1^{p_{12}}) \\ (x_2^{p_{21}} - x_2^{p_{22}}) \\ (x_3^{p_{31}} - x_3^{p_{32}}) \end{bmatrix} = \begin{bmatrix} \sqrt{2} + \sqrt{2}i \\ \sqrt{2} \\ 0 \end{bmatrix}.$$

For  $i = 2, 3$ , (16) is satisfied. The plots of  $I[X_i; Y_i]$  evaluated using Monte-Carlo simulation and  $I[X_i; Y_i]$  evaluated using the high SNR approximation in (14) is shown in Fig. 3. Note that  $R_1$  saturates to a value strictly less than 2 bits/sec/Hz whereas  $R_2$  and  $R_3$  saturates to 2 bits/sec/Hz thus validating Theorem 2.

#### IV. GRADIENT ASCENT BASED ALGORITHM FOR FINITE-SNR

In the previous section, we studied the rate achieved for each transmitter by treating interference as noise at every receiver as  $P$  tends to infinity. In this section, we focus on the finite SNR case. Specifically, the aim is to maximize the sum-rate achieved by treating interference as noise at every receiver with respect to the precoders, i.e.,

$$\max f(V_1, \dots, V_K) = \max \sum_{i=1}^K I[X_i; Y_i] \text{ with } \text{Tr}(V_i V_i^H) \leq 1.$$

This is a non-concave problem in general and difficult to solve. Hence, we propose a gradient-ascent based algorithm to improve the sum-rate starting from some random initialization of precoders. Define the MMSE matrix at Rx- $j$  by

$$E_j = \mathbb{E}[(X - \mathbb{E}[X|Y_j])(X - \mathbb{E}[X|Y_j])^H]$$

where,  $\mathbb{E}$  represents the expectation operator. Define the MMSE matrix at Rx- $j$  with the exclusion of Tx- $j$ 's signal by

$$E_{\mathcal{J}} = \mathbb{E}[(X_{\mathcal{J}} - \mathbb{E}[X_{\mathcal{J}}|Y_j - H_{jj}X_j])(X_{\mathcal{J}} - \mathbb{E}[X_{\mathcal{J}}|Y_j - H_{jj}X_j])^H]$$

The gradient of the sum-rate with respect to the precoder  $V_i$  given by

$$\begin{aligned} \nabla_{V_i} f(V_1, \dots, V_K) &= \nabla_{V_i} \sum_{j=1}^{i=K} I[X_j; Y_j] \\ &= \nabla_{V_i} \sum_{j=1}^{i=K} [I[X_1, X_2, \dots, X_K; Y_j] - I[X_1, X_2, \dots, X_K; Y_j|X_j]] \\ &= \log e \sum_{j=1}^{i=K} H_{ij}^H H_j E_j I_{\sum_{k=1}^{i-1} d_k + 1: \sum_{k=1}^i d_k} \\ &\quad - \log e \sum_{\substack{j=1 \\ j \neq i}}^{i=K} H_{ij}^H H_j E_j I_{\sum_{k=1}^{i-1} d_k - d_j \mathcal{I}(i-j) + 1: \sum_{k=1}^i d_k - d_j \mathcal{I}(i-j)} \end{aligned} \quad (25)$$

where, (25) follows from the relation between the gradient of mutual information and the MMSE matrix obtained in [4]. The matrices

$$I_{\sum_{k=1}^{i-1} d_k + 1: \sum_{k=1}^i d_k} \text{ and } I_{\sum_{k=1}^{i-1} d_k - d_j \mathcal{I}(i-j) + 1: \sum_{k=1}^i d_k - d_j \mathcal{I}(i-j)}$$

select the column numbers from  $\sum_{k=1}^{i-1} d_k + 1$  to  $\sum_{k=1}^i d_k$  of  $E_j$  and  $\sum_{k=1}^{i-1} d_k - d_j \mathcal{I}(i-j) + 1$  to  $\sum_{k=1}^i d_k - d_j \mathcal{I}(i-j)$  of  $E_{\mathcal{J}}$  respectively, where

$$\mathcal{I}(i-j) = \begin{cases} 1 & i > j \\ 0 & i < j. \end{cases}$$

Define  $V = \text{diag}(V_1, V_2, \dots, V_K)$ . The gradient ascent based algorithm for optimizing  $f(V_1, \dots, V_K)$  with respect to the precoders is given in Algorithm 1. During every iteration, whose number is denoted by  $n$ , all the precoders are updated as given in Line

10 of Algorithm 1 where,  $\nabla_V f|_{V=V^{(n-1)}}$  represents  $\text{diag}(\nabla_{V_1} f|_{V_1=V_1^{(n-1)}}, \nabla_{V_2} f|_{V_2=V_2^{(n-1)}}, \dots, \nabla_{V_K} f|_{V_K=V_K^{(n-1)}})$  and  $\nabla_{V_i} f|_{V_i=V_i^{(n-1)}}$  denotes the gradient  $\nabla_{V_i} f$  evaluated at  $V_i = V_i^{(n-1)}$ . If the power constraint for any transmitter Tx- $i$  is violated then,  $V_i^{(n)}$  is projected onto the feasible set with  $\text{Tr}(V_i^{(n)} V_i^{(n)H}) \leq 1$  (see Line 12 of Algorithm 1) [4]. The condition in Line 15 of the algorithm ensures that there is sufficient increase in the objective function. The step size  $t$  of the algorithm is chosen by back-tracking line search with parameters  $\alpha$  and  $\beta$  whose typical values lie between (0.01, 0.3) and (0.1, 0.8) [23]. The proposed algorithm stops when either the number of iterations performed is equal to  $\text{max\_iterations}$  or  $f^{(n-1)} - f^{(n-2)} < \epsilon$  (see Line 5 of Algorithm 1), for some fixed  $\epsilon$ .

---

**Algorithm 1** Gradient Ascent based Algorithm for improving sum-rate

---

```

1: Initialize  $V_i = V_i^{(0)}$  with  $\text{Tr}(V_i V_i^H) \leq 1$ ,  $i = 1, 2, \dots, K$ , and  $t = 1$ .
2: for  $n = 1$  to  $\text{max\_iterations}$  do
3:   Compute  $f^{(n-1)} = f(V_1^{(n-1)}, V_2^{(n-1)}, \dots, V_K^{(n-1)})$ ,
4:      $E_j^{(n-1)}$ , and  $E_j^{(n-1)}$ , for  $j = 1, 2, \dots, K$ .
5:   if  $n > 1$  and  $f^{(n-1)} - f^{(n-2)} < \epsilon$  then
6:     exit for
7:   end if
8:   Compute  $\nabla_V f|_{V=V^{(n-1)}}$ 
9:   do
10:     $V^{(n)} \leftarrow V^{(n-1)} + t \nabla_V f|_{V=V^{(n-1)}}$ .
11:   $V_i^{(n)} \leftarrow \frac{V_i^{(n)}}{\text{Tr}(V_i^{(n)} V_i^{(n)H})}$  if  $\text{Tr}(V_i^{(n)} V_i^{(n)H}) > 1$ , for all  $i$ .
12:  Compute  $f^{(n)} = f(V_1^{(n)}, V_2^{(n)}, \dots, V_K^{(n)})$ .
13:   $t = \beta t$ 
14:  while  $f^{(n)} < f^{(n-1)} + \alpha t \|\nabla_V f|_{V=V^{(n-1)}}\|_F^2$ 
15:     $t = 1$ .
16:  end while
17: end for

```

---

Similar gradient ascent based algorithms have been proposed in the past for optimizing rates in single user MIMO channels [4], [5] and MIMO MAC [8] with finite constellation inputs and precoding. Like in [4] [5] [8], the algorithm does not assume uniform distribution over the elements of the finite constellation. In the following example, we present some simulation results with the proposed algorithm.

*Example 3:* Consider a MIMO GIC with  $K = 3$ ,  $n_{t_i} = n_{r_i} = 2$ ,  $d_i = 1$  for all  $i$ , and the finite constellation used is QPSK.

- 1) *Performance for a fixed channel:* The parameters used in Algorithm 1 are  $\alpha = 0.05$ ,  $\beta = 0.4$ ,  $\epsilon = 0.01$  and  $\text{max\_iterations} = 15$ . The chosen channel matrix is given in (26). The convergence behaviour of Algorithm 1 is shown in Fig. 4 for  $P = -2\text{dB}$ ,  $P = 0\text{dB}$ , and  $P = 2\text{dB}$  with IA precoders of [1] used as initialization. The proposed algorithm terminates well before the  $\text{max\_iterations}$  number for all  $P$  because the condition in Line 5 of Algorithm 1 is satisfied. The sum-rate improvement obtained on application of the proposed algorithm with IA precoder initialization is observed to be 2.35 bits/sec/Hz, 2.12 bits/sec/Hz, and

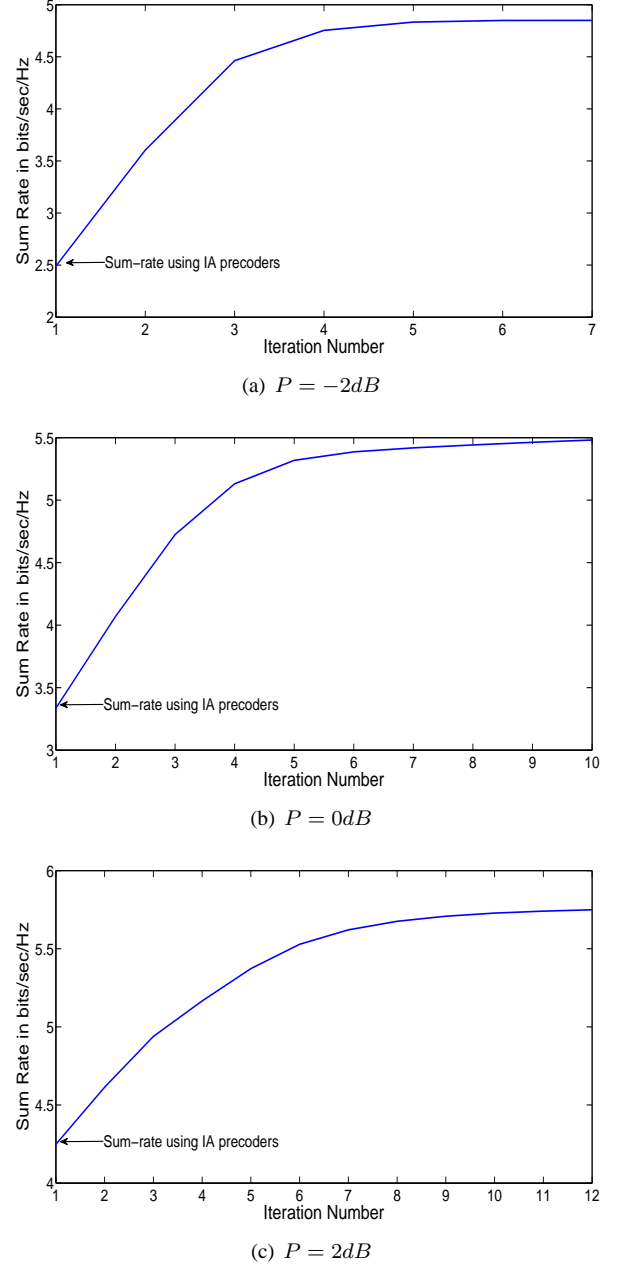


Fig. 4. Sum-rate increase with every iteration for the constant channel case in Example 3.

- 1.5 bits/sec/Hz at  $P = -2\text{dB}$ ,  $P = 0\text{dB}$ , and  $P = 2\text{dB}$  respectively.
- 2) *Performance averaged over channel realizations:* The ergodic sum-rate (i.e., sum-rate averaged with the entries of the channel matrices being taken from  $\mathcal{CN}(0, 1)$ ) obtained by IA precoders of [1] and precoders obtained from Algorithm 1 with the IA precoders as initialization is shown in Fig. 5. The parameters used in Algorithm 1 are  $\alpha = 0.05$ ,  $\beta = 0.4$ ,  $\epsilon = 0.01$ ,  $\text{max\_iterations} = 8$ . The sum-rate increase obtained using the algorithm is 1.07 bits/sec/Hz at  $P = -2\text{dB}$  which then decreases as  $P$  increases, and at  $P = 10\text{dB}$  the sum-rate in-



$$H = \begin{bmatrix} 1.1408 - 0.8637i & 0.1954 - 1.5172i & -0.7038 + 0.3064i & 0.3323 - 0.4278i & 1.3425 - 0.5842i & 0.0766 - 0.7555i \\ 0.9331 + 0.3749i & 1.1588 + 0.1093i & 1.1412 + 0.0561i & 0.2708 + 0.5320i & -1.1639 + 0.3706i & 0.4311 + 1.4726i \\ -0.5206 + 2.7224i & 0.0330 + 1.0577i & -0.1581 - 0.7055i & -1.4823 + 0.7144i & 0.4286 + 0.5008i & -1.0052 + 0.0051i \\ -0.5897 + 0.8322i & -0.9552 + 1.1568i & -0.9685 - 0.7241i & 1.1944 - 0.0556i & -1.3293 - 0.2445i & 0.3540 - 0.7644i \\ 1.3814 - 0.4690i & -0.2731 - 0.6355i & -0.8927 + 0.0044i & -0.1312 + 0.3106i & -1.9229 - 0.2353i & -0.7885 - 0.3229i \\ -0.3043 - 0.2351i & 1.1578 + 0.3676i & 0.9095 - 0.4187i & 0.4161 + 1.2482i & 1.2900 - 0.0196i & 0.4838 + 1.4578i \end{bmatrix}. \quad (26)$$

crease is just 0.1 bits/sec/Hz. The negligible sum-rate improvement at high values of  $P$  is expected because IA precoders are CCSC optimal precoders which saturate the sum-rate at 6 bits/sec/Hz.

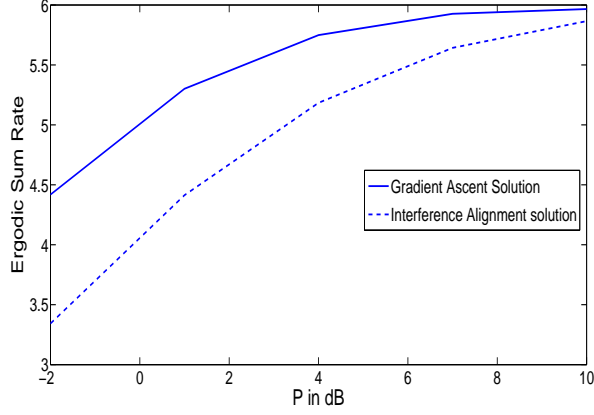


Fig. 5. Ergodic sum-rate in Example 3.

In [1], for sum-DoF analysis zero-forcing matrices are used at every receiver. Here, we have treated interference as noise and have not used any zero-forcing matrices as using them can only reduce the rate  $I[X_i; Y_i]$  because of data-processing inequality [21].

## V. CONCLUSION

The paper discussed linear precoding for  $K$ -user MIMO GIC with finite constellation inputs. We showed that, for constant MIMO GIC with finite constellation inputs, CCSC for every transmitter can be achieved just by using a naive scheme of treating the interference as noise at every receiver, at high SNR. This result is in contrast with the Gaussian alphabet case where, at high SNR, the scheme that treats interference as noise saturates to a value determined by the channel gains for the SISO case. A set of necessary and sufficient conditions for CCSC optimal precoders were derived. It was observed that IA precoders fall under the class of CCSC optimal precoders. However, CCSC optimal precoders are easy to obtain for any given value of channel gains unlike obtaining IA precoders. Finally, a gradient ascent based algorithm was proposed to improve the sum-rate achieved starting from any random initialization of precoders. An example simulation with QPSK input alphabets showed a significant improvement in the ergodic sum-rate achieved by the precoders obtained

from the proposed algorithm compared to the ergodic sum-rate obtained from IA precoders, treating interference as noise at the receivers, at low and moderate SNRs.

## REFERENCES

- [1] V. Cadambe, S. Jafar, "Interference Alignment and Degrees of Freedom for the  $K$  user Interference Channel", IEEE Trans. Info. Theory, Vol. 54, no. 8, pp. 3425-3441, Aug. 2008.
- [2] A. S. Motahari, S. O. Gharan, M. Maddah-Ali, A. K. Khandani, "Real Interference Alignment: Exploiting the Potential of Single Antenna Systems", Available at: arXiv:0908.2282v2 [cs.IT].
- [3] A. Ghasemi, A. S. Motahari, A. K. Khandani, "Interference Alignment for the  $K$  User MIMO Interference Channel", Available at: arXiv:0909.4604v2 [cs.IT].
- [4] D. P. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels", IEEE Trans. Info. Theory, Vol. 52, no. 1, pp. 141-154, Jan. 2006.
- [5] F. Perez-Cruz, M. R. D. Rodrigues, S. Verdu, "MIMO Gaussian Channels With Arbitrary Inputs: Optimal Precoding and Power Allocation", IEEE Trans. Info. Theory, Vol. 56, no. 3, pp. 1070-1084, Mar. 2010.
- [6] C. Xiao, Y. R. Zheng, and Z. Ding "Globally Optimal Linear Precoders for Finite Alphabet Signals Over Complex Vector Gaussian Channels", IEEE Trans Signal Processing, Vol. 59, no. 7, pp. 3301-3314, Jul. 2011.
- [7] J. Harshan and B. Sundar Rajan, "On the Two-User SISO-Gaussian and MIMO-Fading MAC with Finite Input Alphabets", IEEE Transactions on Information Theory, Vol. 57, no. 3, pp.1299-1327, Mar. 2011.
- [8] M. Wang, W. Zeng, and C. Xiao, "Linear Precoding for MIMO Multiple Access Channels with Finite Discrete Inputs", IEEE Trans. Wireless Communications, Vol. 10, no. 11, pp. 3934-3942, Nov. 2011.
- [9] F. Knabe, A. Sezgin, "Achievable Rates in Two-user Interference Channels with Finite Inputs and (Very) Strong Interference", Proc. of IEEE Asilomar SSC 2010, Nov. 7-10, 2010, pp. 2050 - 2054.
- [10] Abhinav Ganesan, B. Sundar Rajan, "Two-User Gaussian Interference Channel with Finite Constellation Input and FDMA", IEEE Trans. Wireless Communications, Vol. 11, no. 7 pp. 2496-2507, Jul. 2012.
- [11] H. Sato, "The Capacity of the Gaussian Interference Under Strong Interference", IEEE Transactions on Information Theory, Vol. 27, no. 6, pp. 786-788, Nov. 1981.
- [12] M. H. M. Costa, A. El Gamal, "The Capacity Region of the Discrete Memoryless Interference Channel with Strong Interference", IEEE Transactions on Information Theory, vol. IT-33, no. 5, pp. 710-711, Sep. 1987.
- [13] X. Shang, G. Kramer, B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels", IEEE Trans. Inf. Theory, vol. 55, no. 2, pp. 689-699, Feb. 2009.
- [14] A. Motahari, A. Khandani, "Capacity bounds for the Gaussian interference channel", IEEE Trans. Inf. Theory, Vol. 55, no. 2, pp. 620-643, Feb. 2009.
- [15] V. Annapureddy, V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region", IEEE Trans. Inf. Theory, Vol. 55, no. 7, pp. 3032-3050, Jul. 2009.
- [16] X. Shang, G. Kramer, B. Chen, "Throughput Optimization in Multiuser Interference Channels", Proc. of IEEE MILCOM 2008, Nov. 16-19, 2008, pp. 1-7.
- [17] X. Shang, B. Chen, G. Kramer, H. Vincent Poor, "Capacity Regions and Sum-Rate Capacities of Vector Gaussian Interference Channels", IEEE Trans. Inf. Theory, Vol. 56, no. 10, pp. 5030-5044, Oct. 2010.
- [18] V. Annapureddy, V. Veeravalli, "Sum Capacity of MIMO Interference Channels in the Low Interference Regime", IEEE Trans. Inf. Theory, Vol. 57, no. 5, pp. 2565-2581, May 2011.
- [19] A. El Gamal and Young-Han Kim, "Network Information Theory", Cambridge University Press, 2012.

- [20] M. Razaviyayn, M. Sanjabi, Z. Luo, "Linear Transceiver Design for Interference Alignment: Complexity and Computation", IEEE Trans. Inf. Theory, Vol. 58, no. 5, pp. 2896-2910, May 2012
- [21] T. M. Cover and J. A. Thomas, "Elements of Information Theory", second edition - Wiley Series in Telecommunications, 2004.
- [22] O. Gonzalez, I. Santamara and C. Beltran, "A general test to check the feasibility of linear interference alignment", IEEE ISIT 2012, Cambridge, Jul. 2012, pp. 2491-2495.
- [23] S. P. Boyd and L. Vandenberghe, "Convex Optimization". Cambridge University Press, 2004.